UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/851,066	05/07/2001	Tongwei Liu	HP-10012392	2859

7590 04/17/2006

HEWLETT-PACKARD COMPANY
Intellectual Property Administration
P.O. Box 272400
Fort Collins, CO 80527-2400

EXAMINER	
LE, MIRANDA	
ART UNIT	**PAPER NUMBER**
2167	

DATE MAILED: 04/17/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/851,066	LIU ET AL.	
	Examiner	Art Unit	
	Miranda Le	2167	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *01 February 2006*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-22* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-22* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

1. Claims 1-22 are pending in this application. Claims 1, 8, 15, 22 are independent claims. In the Amendment, no claims have been amended. This action is made Final.

Claim Rejections - 35 USC § 101

2. 35 U.S.C. § 101 reads as follows:

> "Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter or any new and useful improvement thereof, may obtain a patent therefore, subject to the conditions and requirements of this title".

3. Claims 1, 22 are rejected under 35 U.S.C. § 101 because the claimed invention is directed to non-statutory subject matter.

Claim 1 fails to provide a practical application that produces a useful, concrete and tangible result. Specifically, claim 1 lacks a tangible result. A tangible result, in accordance with the current "Interim Guidelines for Examination of Patent Applications for Patent Subject Matter Eligibility" is a real world result. In other words, the result is required to be outside of the "abstract". The method of claim 1 defines a sequence of operational steps that encompasses within its scope merely a set of mental manipulations that provide an output remaining in the mental realm. As such, the "result" of the method of claim 1 resides in the "abstract" and is, therefore, not a "real world" result.

Further, the claim appears to have no claimed result under the condition where the record does not have information for all of the variables but the first classification tool does not need the missing information. In this case, the claim appears to produce no result, which would be fixed by including a limitation directed towards the first tool being used when information is missing but not required.

Claim 22 has the same type of issues as claim 1 therefore, is rejected under similar rationale.

Claim Rejections - 35 USC § 102

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless:
>
> (e) the invention was described in
> (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or
>
> (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

5. Claims 1-6, 8-13, 15-20, 22 are rejected under 35 U.S.C. 102(e) as being anticipated by Hadzikadic et al. (US Pub. No. US 2002/0059202).

Hadzikadic anticipated independent claims 1, 8, 15, 22 by the following:

As to claims 1, 8, 15, Hadzikadic teaches a method comprising: receiving a record comprising a plurality of variables (i.e. attributes, [0077]), wherein said record (i.e. an instance, [0077]) comprises information for at least some of said variables ([0028, 0031, 0077, 0126, 0127]);

if said record comprises information for all of said variables (i.e. using all attributes, [00126]), using said information with a first classification tool adapted to classify said record

(i.e. a first classification tree from the data set of instances to be classified, using all attributes, [00126]) ([0028, 0031, 0077, 0126, 0127]); and

if said record comprises information for some, but not all (i.e. less attributes, [0127]), of said variables, using said information with a second classification tool (i.e. subsequent classification tree, [0127]) instead of with said first classification tool to classify said record in response to determining that said first classification tool requires a particular item of information that is missing (i.e. incomplete instances in [0034], or with instances containing fewer attributes, in [0129]) from said information ([0028, 0031, 0034, 0077, 0126, 0127, 0129]).

As per claim 22, Hadzikadic teaches: a method for classifying an information record that comprises a plurality of variables, said method comprising:

ranking said plurality of variables according to their respective influence on said classifying (i.e. weighted attribute, [0033]);

grouping said plurality of variables into subsets of variables using said ranking, wherein a classification tree is computed for each of said subsets (i.e. using weighted attribute relevance in the calculation of similarity and cohesiveness value in order to build an unbiased classification tree, [0033], [0086]);

receiving a record comprising information for at least some of said variables associated with record and, if said record comprises information for all of said associated variables (i.e. using all attributes, [00126]), using said information with a first classification tree adapted to classify said record, wherein said first classification tree is based on a substantially complete set of information for said plurality of variables (i.e. a first classification tree from the data set of

instances to be classified, using all attributes, [00126]) ([0028, 0031, 0074, 0076, 0077, 0126, 0127]); and

if said record comprises information for some , but not all, (i.e. less attributes, [0127]),of said associated variables, using said information with a second classification tree (i.e. subsequent classification tree, [0127]) instead of with said first classification tree to classify said record when said first classification tree requires a particular item of information that is missing (i.e. incomplete instances in [0034], or with instances containing fewer attributes, in [0129]) from said information, wherein said second classification tree is based on information for one of said subsets of variables, wherein said one of said subsets does not include said particular item of information that is missing (i.e. Computational server 135 retains at least one classification tree which uses less attributes than the first classification tree, [0127]) ([0028, 0031, 0034, 0077, 0126, 0127, 0129]).

As to claims 2, 9, 16, Hadzikadic teaches first classification tool and said second classification tool are a first classification tree and a second classification tree, respectively ([0078], [0079], [0080]).

As to claims 3, 10, 17, Hadzikadic teaches first classification tree is computed using a substantially complete set of information for said plurality of variables (i.e. builds a first classification tree from the data set of instances to be classified, using all attributes, and determines the prediction accuracy for the first classification tree, [0126]) and wherein said second classification tree is computed using information for a subset of said plurality of

variables, wherein said subset does not include said particular item of information that is missing (i.e. builds subsequent classification trees, using fewer attributes for each instances to be classified with each subsequent build, e.g. missing information, determines the prediction accuracy of each subsequent classification trees, discarding those classification trees that do not meet the prediction accuracy acceptability criteria, retains at least one classification tree which uses less attributes than the first classification tree and which predicts with results that fall within the prediction accuracy acceptability criteria, [0127], [0037], [0042]).

As to claims 4, 11, 18, Hadzikadic teaches "ranking said plurality of variables according to their respective influence on said classifying" at [0086];

"grouping said plurality of variables into subsets of variables using said ranking" at [0086].

As to claims 5, 12, 19, Hadzikadic teaches computing a classification tree for each one of said subsets ([0127], [0078], [0079], [0080], [0042]).

As to claims 6, 13, 20, Hadzikadic teaches "said record comprises customer information for a client, wherein content is selected for delivery to a customer according to said classifying of said record" at [0043], [0044].

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

7. Claims 7, 14, 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Hadzikadic et al. (US Pub. No. US 2002/0059202), in view of Gehrke et al. (Tutorial notes of the fifth ACM SIGKDD internation conference on Knowledge discovery and data mining).

As to claims 7, 14, 21, Hadzikadic does not specifically teach "substituting a default value for said particular item of information that is missing". However, Gehrke teaches this limitation at page 32 (missing values handled by surrogate splits).

Thus, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the teachings of Hadzikadic with the teachings of Gehrke to include "substituting a default value for said particular item of information that is missing" in order to efficiently construct predictor trees from very large training databases by utilizing tree pruning and missing value handling methods.

Response to Arguments

8. Applicant's arguments filed 02/01/2006 have been fully considered but they are not persuasive.

Applicant argues that:

(a) The 101 rejection is improper.

(b) Hadzikadic does not teach the use of a classification tree to classify a record.

(c) Hadzikadic does not teach the instance of information to be classified is incomplete.

The Examiner respectfully disagrees for the following reasons:

Per (a), regarding applicant's argument of paragraph 2 (page 7), it is agreed that, currently, "mental step or human step tests" are not to be applied. Nonetheless, in this Office Action, the "practical application test" is applied.

It is further noted that applicant's arguments alluding to support of his position in State StreetBank & Trust Co. v. Signature Financial, 149 F.3d 1368, 1373 (Fed. Cir. 1998) are merely conclusory and are not supportable under State Street with regard to applicant's specific claims.

Per (b), applicant's argument regarding examiner relies on Hadzikadic that was used two years ago, has been fully considered but it is believed that the claim language as presented is still read on by the Hadzidadic reference at the cited paragraph in the claim rejections.

Applicant's invention and Hadzidadic's are directed to the same method for classifying sets of data and predicting new data values or incomplete data according to the following reasons:

1. Applicants invention directs to a method and	Hadzikadic's invention also directs to a

system for accurately predicting the class membership of a record wherein information for one or more of the variables in the record is missing (Specification, page 4, lines 8-10).	method and system for predicting the class as *"The prediction coverage designates which classification structures will be queried for a prediction"*([0031]). Hadzikadic's invention also directs to a method and system for predicting the class as *"The prediction coverage designates which classification structures will be queried for a prediction"*([0031]). Hadzikadic's also direct to the method for classifying a record wherein information for one or more of the variables in the record is missing as disclosed in [0034 *"incomplete instances"*, or in [0129] *"with instances containing fewer attributes"*.
2. Applicants invention discloses multiple classification tools (e.g. classification trees or partition trees) that are generated from an training data set that contains little or missing data and where the class assignments are known (Specification, page 4, lines 10-13).	Hadzikadic's invention discloses the multiple classification tools (i.e. multiple classification structures, [0028]), which are built from a training data set (i.e. Such a scenario is used to construct a classification tree from training data that is input manually or read in as a file, [0074]).

3. A substantially complete set of training data is used to compute a first classification tree (Specification, page 4, lines 13-15).	Hadzikadic's invention discloses a first classification tree applied to a substantially complete set of training data as *"Computational server 135 builds a first classification tree from the data set of instances to be classified, using all attributes" at [0126].*
4. Subsets of the variable in the training data are selected and used to compute other classification trees (Specification, page, lines 15-16).	Hadzikadic's discloses subsets of variable (i.e. uses less attributes, in Para [0127]) which are used to build other classification trees (i.e. subsequent classification trees, in Para [0127]) as in paragraph [0127] bellow: *"Next, computational server 135 builds subsequent classification trees, using the classification process previously discussed in FIG. 2 through FIG. 5. <u>However, computational server 135 includes fewer attributes for each instance to be classified with each subsequent build</u>. Again, computational server 135 determines the prediction accuracy of each subsequent*

	classification tree, discarding as unacceptable, those classification trees that do not meet the prediction accuracy acceptability criteria. Computational server 135 retains at least one classification tree which uses less attributes than the first classification tree and which predicts with results that fall within the prediction accuracy acceptability criteria".
5. Variable are selected for inclusion in a subset based on how strongly they influence the prediction of class membership (Specification, page 4, lines 16-18).	Hadzikadic's discloses prediction of class membership as in Para [0021] below: *"The present invention also provides methods for predicting the attribute values of instances (i.e., data). <u>One embodiment of such methods receives a pre-existing classification structure and an instance to be predicted, determines the best host for the instance to be predicted, and optionally places the instance to be predicted in the classification structure</u>. The instance to be predicted, in fact, comprises at least one attribute value to be predicted. The best host determination and the optional placement of*

	the instance to be predicted processes are performed for prediction using the same techniques as for classification. Prediction may further involve determining a confidence level, or prediction weight, for an occurrence of at least one possible value for the attributes to be predicted".

Firstly, Hadzikadic teaches the two steps of building a classification tree and classifying a record (i.e. more data to classify, step 209) as disclosed in Fig. 2 and in paragraph [0081]:

Once a best host determination is made, a delta cohesiveness measure (.DELTA.CM) is calculated for I and each child of the best host in step 207. The delta cohesiveness measure indicates the degree of cohesion between I and each best host child. This step is explained in further detail in FIG. 4. The instance I is then inserted into the classification tree in step 208. The rules for I's placement in relation to the best host children will be presented in detail in the discussion of FIG. 5. In step 209, it is determined whether more data is to be classified. If more data is to be classified, then the next instance is received in step 201, and classification continues in steps 202 through 209. If no more data is to be classified, then classification ends at step 220, resulting in a tree structure of classified instances (i.e., data points).

Hadzikadic teaches the multiple classification trees for classify a record in paragaph [0028] as recited below:

Another embodiment of the present invention facilitates the building of a federated group of multiple classification structures from the same data set of instances to be classified (or from multiple data sets, which are treated as one large data set), by receiving a data set (or data sets)

comprising at least one instance to be classified, receiving a federated group number and then constructing a group of classification structures or other learning methods from the same data set (or sets) of instances to be classified. Federation is not limited to classification structures built by the various embodiments of the present invention. Federation may also be applied to build and utilize classification structures based upon other learning methods.

Secondly, Hadzikadic teaches the use of one classification instead (i.e. designates which classification structures) of another is disclosed in paragraph [0076]:

Embodiments of the present invention may use various prediction profiles. For example, one embodiment of the present invention provides for predicting by receiving a group of classification structures, receiving an instance with at least one attribute to predict, using each classification structure that is a predicting member of a subset of a group of classification structures to predict on the instance with at least one attribute to predict, and coalescing (or combining) the prediction profiles from each predicting member. The prediction coverage designates which classification structures will be queried for a prediction. As another example, an embodiment provides an Ameliorate prediction profile. The Ameliorate prediction profile is the average of the prediction profiles returned from each queried classification structure.

Thirdly, Hadzikadic teaches the fist classification tool (i.e. a first classification tree) to classify the data including all attributes (i.e. using all attributes) in paragraph [0126] recited below:

To illustrate how classification and prediction go hand-in-hand, note that an embodiment of the present invention allows an operator to determine the importance of attributes, and reconstitute classification trees to decrease the number of attributes per node that must be processed during the prediction process. Computational server 135 receives from secondary storage 125 a data set comprising instances to be classified, wherein these instances to be classified are comprised of at least two attributes and corresponding relevance values. Computational server 135 builds a first classification tree from the data set of instances to be

classified, using all attributes, and determines the prediction accuracy for the first classification tree. In one embodiment, the prediction accuracy is determined as the average performance of the model on a test partition, wherein the test partition comprises a training or standard data set. Computational server 135 establishes at least one prediction accuracy acceptability criterion. In another embodiment, a prediction accuracy acceptability criterion is a configuration parameter, such as a threshold value for a confidence level. A threshold value may be provided by the system provider, a system administrator, the user, or automatic based on specified attribute.

Fourthly, Hadzikadic teaches the second classification tool (i.e. subsequent classification trees) to classify the data which comprises information for some, but not all of said variables (i.e. less attributes, in paragraph [0127]), or incomplete information (i.e. incomplete instances, in paragraph [0034]) recited below:

Next, computational server 135 builds subsequent classification trees, using the classification process previously discussed in FIG. 2 through FIG. 5. However, computational server 135 includes fewer attributes for each instance to be classified with each subsequent build. Again, computational server 135 determines the prediction accuracy of each subsequent classification tree, discarding as unacceptable, those classification trees that do not meet the prediction accuracy acceptability criteria. Computational server 135 retains at least one classification tree which uses less attributes than the first classification tree and which predicts with results that fall within the prediction accuracy acceptability criteria.

Other advantages include the efficient and accurate prediction of attribute profiles for new and incomplete instances. An embodiment of the present invention uses weighted attribute relevance, calculates similarity and delta cohesiveness measurement values to find the best host in a classification tree, and uses sum pair calculations to reduce the number of nodes processed, thereby increasing prediction accuracy and efficiency.

It is thus clearly shown that Hadzikadic does teach the use of a classification tree to classify a record and/or applying classification tools (trees) to classify a record.

Per (c), Hadzikadic teaches the instance of information to be classified is incomplete (i.e. less attribute, in paragraph [00127], or fewer attributes in paragraph [0127, 0129]).

The alternate method, which is less time intensive, is to allow an embodiment of the present invention to determine prediction profiles over all of the attributes. Then the system can automatically, or the operator can manually, analyze those profiles for attribute values with confidence level values below a pre-specified threshold level. The system, automatically, or the operator, manually, can then remove those attributes deemed unimportant and rebuild the classification tree with instances containing fewer attributes. In one embodiment, this process may be performed several times until the average confidence level over all of the instance attribute values drops below a pre-specified acceptance level. The resultant classification tree is that classification tree which contains nodes with the fewest instance attributes and still meets the pre-specified acceptance level.

In addition, Hadzikadic further discloses the incomplete information (i.e. incomplete instances) in paragraph [0034] as recited below:

Other advantages include the efficient and accurate prediction of attribute profiles for new and incomplete instances. An embodiment of the present invention uses weighted attribute relevance, calculates similarity and delta cohesiveness measurement values to find the best host in a classification tree, and uses sum pair calculations to reduce the number of nodes processed, thereby increasing prediction accuracy and efficiency.

For the above reasons, it is evident that the Hadzikadic system cannot be distinguished from the claim invention. Arguments as raised are moot since all claim limitations relevant to this issue have been addressed accordingly.

Conclusion

9. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Miranda Le whose telephone number is (571) 272-4112. The examiner can normally be reached on Monday through Friday from 8:30 AM to 5:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Jean Homere, Esq., can be reached on (571) 272-3780. The fax number to this Art Unit is 571-273-8300.

Any inquiry of a general nature or relating to the status of this application should be directed to the Group receptionist whose telephone number is (703) 305-3900.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Miranda Le
April 13, 2006

SEVA A. ROBINSON
PRIMARY EXAMINER